Feasibility of MLOps-based healthcare pipelines in ensuring the Cybersecurity Framework

Antonio Robustelli[1,*], Alberto Marfoglia[1], Christian D'Errico[1], Sabato Mellone[2] and Antonella Carbonaro[1]

[1]Department of Computer Science and Engineering, University of Bologna, Italy
[2]Department of Electrical, Electronic and Information Engineering "Guglielmo Marconi", University of Bologna, Italy

Abstract

The recent advances in Artificial Intelligence (AI) are radically transforming the healthcare sector. Implementing the related solutions presents significant challenges, ranging from managing data quality and heterogeneity to compliance with stringent regulations (e.g., GDPR and HIPAA). In this context, MLOps emerges as a crucial solution to address these issues through a set of practices and tools. As a result, MLOps-based pipelines play a pivotal role in the effective management of Machine Learning (ML) models, which is vital to support diagnostic and prognostic activities. On the other hand, the development of healthcare systems should also consider several cybersecurity aspects required by the same regulations. To this end, the Cybersecurity Framework (CSF) 2.0, developed by the National Institute of Standards and Technology (NIST), describes updated guidelines to mitigate cybersecurity risks. Therefore, adopting MLOps with the support of the CSF represents an essential step for enabling the transition of ML models to enabled devices and improving the security of healthcare systems. For this reason, in this work, we present the high-level architecture of an MLOps pipeline employed by the DARE (DigitAl lifelong pRevEntion) foundation. Moreover, we also analyze its feasibility in satisfying CSF requirements, with particular emphasis on those related to data security, detection, and recovery.

Keywords

MLOps pipeline, CSF, Healthcare, Machine Learning

1. Introduction

In recent years, Artificial Intelligence (AI) and Machine Learning (ML) have revolutionized the healthcare sector, providing powerful tools to face complex challenges. For instance, many ML-based models were increasingly experimented to assist physicians in a wide range of activities, such as disease diagnosis [1, 2], treatment personalization [3, 4], and patient monitoring [5, 6]. However, despite the excellent results obtained, most of the attempts to employ ML-based approaches have not overcome the prototypical status [7]. This generally happens because the transition of ML prototypes to ML-enabled medical devices represents a complex process due to the numerous and strict existing regulations (e.g., GDPR and HIPAA) [8]. Moreover, this transition requires a complex interdisciplinary endeavour in which data scientists need to collaborate with software engineers, operations teams, domain experts, and end users to build a successful product [9].

For this reason, new ML engineering practices, known under the terminology of Machine Learning Operations (MLOps), are emerging to support this transition [10, 11]. Consistent with the principles of DevOps (Development Operations), MLOps aims to bring automation to the development workflow of ML-enabled systems by streamlining the ML models' lifecycle [12, 13]. To this end, MLOps can enhance operational efficiency, allowing teams to focus on innovation and strategic goals rather than repetitive tasks [14, 15]. Furthermore, the MLOps scalability enables organizations to manage large datasets and release models more easily [16]. Consequently, MLOps-based pipelines have gained a

MLOps25: Workshop on Machine Learning Operations. October 25, 2025, Bologna, Italy
*Corresponding author.
antonio.robustelli2@unibo.it (A. Robustelli); alberto.marfoglia2@unibo.it (A. Marfoglia); christian.derrico2@unibo.it (C. D'Errico); sabato.mellone@unibo.it (S. Mellone); antonella.carbonaro@unibo.it (A. Carbonaro)
0000-0003-3423-3374 (A. Robustelli); 0009-0000-5857-2376 (A. Marfoglia); 0009-0002-5784-0492 (C. D'Errico); 0000-0001-7688-0188 (S. Mellone); 0000-0002-3890-4852 (A. Carbonaro)

strong interest in the healthcare industry, in which the correct management of the models' lifecycle is essential to support diagnostic and prognostic activities [17, 18].

However, despite being crucial in healthcare, only some cybersecurity aspects have been relatively investigated [19]. Instead, a further effort should concern the exploitation of MLOps to make the related development environments more compliant with the most notable security frameworks. Moreover, the increasing complexity and spread of cyber threats force organizations to adopt even more advanced mitigation strategies. For instance, model inversion attacks represent one of the pressing threats due to their ability to exploit model outputs and, through several reverse engineering steps, reconstruct training data or make inferences on them [20, 21]. Consequently, the deployment of ML models also requires reliable mechanisms of Role-Based Access Control (RBAC), in which each user can access only to specific data or artifacts, and be logged during the model's lifecycle [22].

In response to these challenges, the Cybersecurity Framework (CSF) 2.0, defined by the National Institute of Standards and Technology (NIST), provides updated guidelines to address security challenges in an ever-evolving technological landscape [23]. The CSF is a strategic tool to protect digital assets, enhance stakeholder trust, and improve the organizational's resilience. Since each organization presents unique risks, varying risk tolerances, specific missions, and desired objectives, the CSF does not embrace a one-size-fits-all approach. Instead, it recommends its implementation by employing several emerging technologies and solutions [23].

Therefore, adopting MLOps with the support of the CSF represents an essential investment to help the transition of ML-based models to enabled devices. Moreover, their employment also improves the security of healthcare organizations like DARE (DigitAl lifelong pRevEntion) [24], a foundation financed by the Italian Ministry for University & Research (MUR) to foster collaboration between healthcare, academia, industry, and policymakers. In detail, DARE aims to become a national reference for digital prevention technologies, enhance health promotion, and enable lifelong prevention. To achieve such goals, DARE needs a compliant infrastructure capable of hosting different research studies, managing healthcare data securely, and developing reliable AI models.

For this reason, in order to study the feasibility of MLOps pipelines in ensuring several cybersecurity aspects, we first define a high-level MLOps pipeline employed by the DARE foundation. Then, by adopting the CSF, we analyze its validity in ensuring different requirements, with particular emphasis on those related to data security, detection, and recovery.

The main contributions of this work can be summarized as follows:

1. We define the high-level architecture of an MLOps pipeline employed in a healthcare foundation;
2. We adopt the CSF to analyze the pipeline's feasibility in ensuring different requirements, namely data security, detection, and recovery.

The remainder of the paper is organized as follows. Sec. 2 will present the related works on MLOps pipelines employed in healthcare scenarios. Sec. 3 will report an overview of the CSF's structure and MLOps. Then, Sec. 4 will define the architecture of our MLOps pipeline employed for the DARE foundation. Finally, Sec. 5 will analyze the feasibility of the pipeline in ensuring CSF requirements, while Sec. 6 will present the conclusions and future work.

2. Related Works

Several studies have explored the application of MLOps frameworks in healthcare with the aim of enhancing the development, deployment, and management of ML models in clinical settings [25, 10, 11]. These contributions are essential for bridging the gap between prototypical research and practical implementation in such domains, which are typically highly regulated [8, 26]. MLOps practices can thus provide crucial benefits such as reproducibility, maintainability, trackability, and regulatory compliance [12, 13].

To this end, A. Basile et al. [27] have proposed a comprehensive MLOps pipeline by integrating many famous tools for version control, experiment tracking, and continuous monitoring. Instead, V.

Moskalenko et al. [17] have introduced several practices designed to enhance the robustness of medical diagnostic systems. In detail, they have implemented additional pipeline stages to face prevalent issues related to healthcare environments, such as the risks associated with adversarial attacks, fault injections, and distribution shifts.

Moreover, an additional effort has been made by implementing several MLOps-based tools, such as that proposed by A. Krishnan et al. [28]. In detail, they have developed CyclOps, an open-source framework to address the fragmented nature of ML tools in healthcare units. The achieved results, derived by predicting in-hospital and decompensation mortality, have proven the effectiveness of CyclOps in ensuring the adaptability, scalability, and reliability of the developed ML models. Similar outcomes have been shown by Advanced Notebook (ADVN), a tool proposed by G. Danciu et al. [29] to standardize data ingestion and manage ML models in two major EU projects: iHELP and RETENTION. In such studies, the authors have employed ADVN to predict the risk of pancreatic cancer using urinary biomarkers (in iHELP) and estimate heart failure survival (in RETENTION). The related use cases have highlighted the versatility of ADVM in handling various medical challenges and improving the development process. Finally, T. Granlund et al. [30] have employed Oravizio, a CE-certified software used in joint replacement surgery risk assessments, to demonstrate how MLOps can ensure data privacy laws and regulatory standards without compromising performance. To accomplish this, the authors have designed a continuous training pipeline to automate data validation, model re-training, and the generation of regulatory-compliant reports.

However, as shown in Tab. 1, only some cybersecurity aspects have been relatively investigated [19]. Instead, since the adoption of CSF with MLOps represents an essential investment to support the transition of ML-based models to enabled devices, this study aims to analyze the feasibility of MLOps healthcare pipelines in ensuring several cybersecurity requirements.

Reference	Feature	Gap & Novelty
V. Moskalenko et al. [17]	– Enhance the robustness of medical diagnostic systems	✗ No use of well-known cybersecurity techniques or protocols ✓ Analyze risks like adversarial attacks, fault injections, and distribution shifts
G. Danciu et al. [29]	– Integration of ADVN to manage data ingestion and ML models	✗ No cybersecurity analysis is done (just some considerations) ✓ Employ anonymized data
T. Granlund et al. [30]	– Continuous generation of regulatory-compliant reports	✗ No cybersecurity analysis is done (just some considerations) ✓ Ensure data privacy laws and regulatory standards
Our proposal	– Adoption of the CSF 2.0	✓ Feasibility analysis of data security, detection, and recovery

Table 1
Security-oriented comparison between the existing MLOps-based solutions and our proposal. For each Reference, the following table highlights the related Feature (–), Gap (✗), and Novelty (✓).

3. Background

This section provides an overview of the fundamental concepts related to the proposed pipeline. For this reason, we first recall the structure of CSF and the related Functions. Then, we briefly summarize the main characteristics of MLOps, which represent the adopted framework.

3.1. The Cybersecurity Framework 2.0

The Cybersecurity Framework (CSF) 2.0, defined by NIST, is designed to help organizations of all sizes and sectors to manage and reduce their cybersecurity risks [23]. Since each organization has different risks and desired objectives, the CSF does not embrace a one-size-fits-all approach. Instead, the way how organizations implement CSF can vary and involve different emerging solutions. For this reason, the CSF describes the cybersecurity outcomes and requirements for a broad audience, including executives, managers, and practitioners, regardless of their cybersecurity expertise. As shown in Fig. 1, such outcomes are mapped into a dedicated list known as Core, and which consists of 6 Functions, namely Govern (GV), Identify (ID), Protect (PR), Detect (DT), Respond (RS), and Recover (RC).

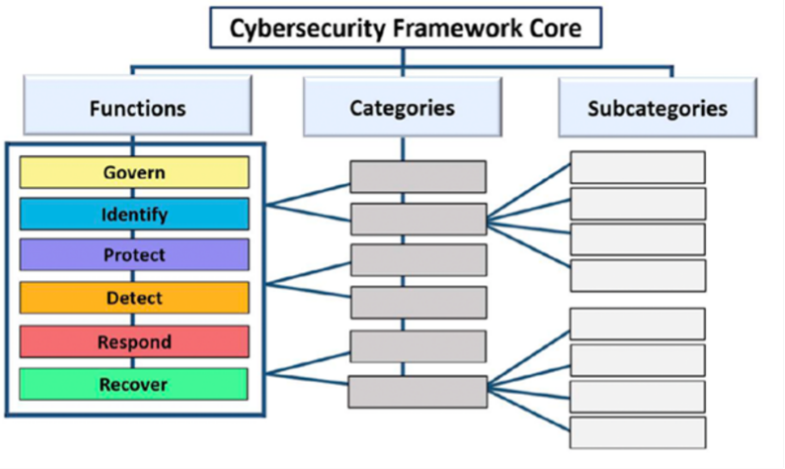

Figure 1: The high-level Core's structure [23].

These outcomes do not represent a checklist of actions to perform but, instead, high-level requirements that an organization should ensure in relationship with its use cases. In detail, each Function is divided into Categories that represent a subset of cybersecurity outcomes. Finally, subcategories further divide each Category into more specific outcomes. Fig. 2 reports the Categories associated with each Function and the related identifiers.

Function	Category	Category Identifier
Govern (GV)	Organizational Context	GV.OC
	Risk Management Strategy	GV.RM
	Roles, Responsibilities, and Authorities	GV.RR
	Policy	GV.PO
	Oversight	GV.OV
	Cybersecurity Supply Chain Risk Management	GV.SC
Identify (ID)	Asset Management	ID.AM
	Risk Assessment	ID.RA
	Improvement	ID.IM
Protect (PR)	Identity Management, Authentication, and Access Control	PR.AA
	Awareness and Training	PR.AT
	Data Security	PR.DS
	Platform Security	PR.PS
	Technology Infrastructure Resilience	PR.IR
Detect (DE)	Continuous Monitoring	DE.CM
	Adverse Event Analysis	DE.AE
Respond (RS)	Incident Management	RS.MA
	Incident Analysis	RS.AN
	Incident Response Reporting and Communication	RS.CO
	Incident Mitigation	RS.MI
Recover (RC)	Incident Recovery Plan Execution	RC.RP
	Incident Recovery Communication	RC.CO

Figure 2: The CSF Categories with the related identifiers [23].

3.2. Machine Learning Operations

The acronym MLOps (Machine Learning Operations) represents the evolution of DevOps (Development Operations) practices applied to the lifecycle of ML models [12, 15]. MLOps offers significant advantages by streamlining and automating the complex lifecycle of ML models [12, 13]. For instance, MLOps enhances operational efficiency, allowing teams to focus on innovation and strategic goals rather than repetitive tasks [14, 15]. MLOps ensures that models are continuously updated, tested, and monitored for optimal performance [14]. Moreover, it minimizes risks, performance degradation, and data drift [13], enabling organizations to manage large datasets and released models [16].

Therefore, MLOps focuses on all aspects of ML models, from the requirement analysis to monitoring in production. In detail, the lifecycle involves several stages, each closely tied to monitoring, maintenance, and continuous updates. In MLOps, the lifecycle does not end with the initial training phase but extends to ongoing management and optimization, enhancing the model's ability to adapt to dynamic changes. Also, the employment of a Continuous Integration and Continuous Delivery (CI/CD) approach significantly contributes to the model's stability and reliability over time [31].

However, due to the growth of cybersecurity threats, it has become essential for MLOps to manage security aspects. In such cases, the adopted terminology is known as SecMLOps or MLSecOps [22]. Although it is difficult to identify a common and widely accepted definition, we can refer to that provided by B. Ghosh [32]. In detail, he defined MLSecOps as "implementing and managing a set of processes, tools, and best practices that are designed to secure machine learning models and the systems that support them. It aims to address the unique challenges of securing ML models at scale."

4. The proposed MLOps pipeline

As previously mentioned, the integration of MLOps represents a crucial step towards adopting safe, reliable, and effective ML-based approaches, which are increasingly experimented in the healthcare domain to assist physicians in a wide range of activities, such as disease diagnosis [1, 2], treatment personalization [3, 4], and patient monitoring [5, 6]. However, despite the excellent results, these approaches require rigorous management to ensure data quality, model robustness, and compliance with privacy and security regulations [19]. For this reason, we first present the high-level architecture of a pipeline employed in a real healthcare foundation. Then, we introduce the MLOps development cycle inside our pipeline by mapping the related steps. For clarity, since it would be out of the scope of this work, the provided definition refers only to a high-level architecture. Consequently, we do not report a technical definition of the pipeline, but remand to [17, 27, 33] for more detailed implementations.

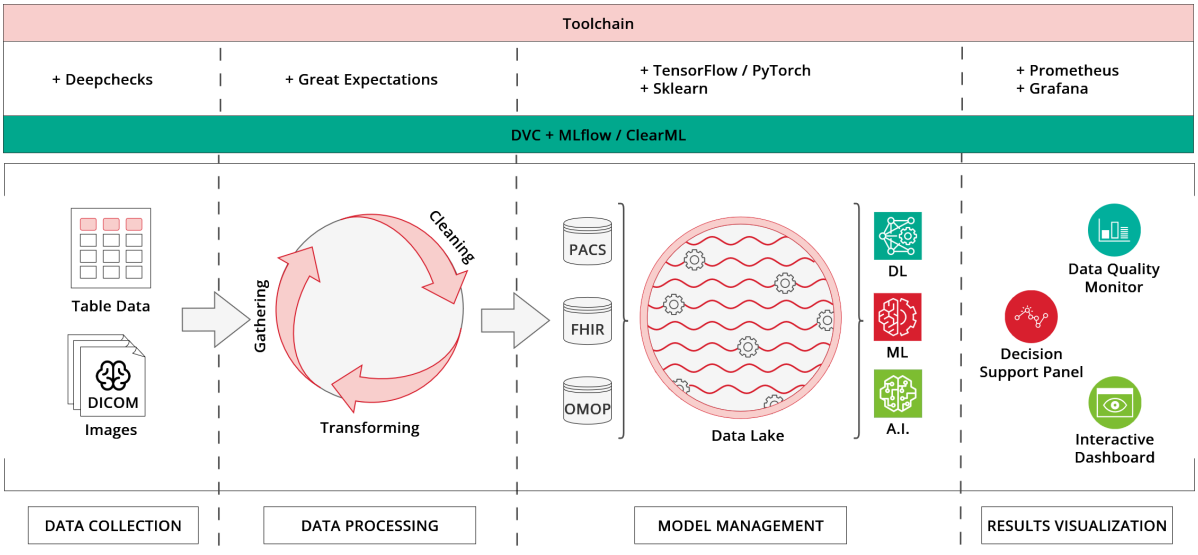

Figure 3: The High-level pipeline architecture and the related toolchain.

4.1. High-Level Architecture

The proposed pipeline aims to securely manage healthcare data and deploy reliable AI models while ensuring adherence to regulatory and ethical standards. Therefore, intending to provide a compliant infrastructure capable of hosting different research studies for the DARE foundation [24], we structured our pipeline into the following steps shown in Fig. 3:

1. **Data Collection:** data are systematically gathered from multiple sources, including table data, DICOM medical imaging files, and other structured or unstructured datasets. Furthermore, this step incorporates rigorous provenance tracking to ensure compliance with consent protocols and legal approvals before data ingestion;
2. **Data Processing:** once gathered, data undergoes comprehensive processing activities to ensure its suitability for the research goals. In detail, this step involves systematic cleaning and transformation procedures to enhance the data quality and consistency, addressing also issues related to missing values and inaccuracies. The processed data is thus standardized and aligned with established clinical frameworks, including Fast Healthcare Interoperability Resources (FHIR) and Observational Medical Outcomes Partnership (OMOP), to facilitate seamless interoperability;
3. **Model Management:** next, the processed data are pèriodically organized and stored within a data lake infrastructure, which is capable of handling several data formats. To this end, the Model Management integrates relational databases, Picture Archiving and Communication Systems (PACS), and servers that adhere to widely recognized clinical standards. This step also incorporates mechanisms for controlled data retention, avoiding unnecessary storage prolongation. However, the primary aim of this step is the development of AI and ML models. For this reason, these models leverage the processed data to extract relevant insights for the decision-making activities;
4. **Results Visualization:** finally, the developed models are deployed inside real healthcare applications to face different tasks. For this reason, it is crucial to monitor the related performance by tracking each execution along with the associated input configurations. Consequently, the pipeline must also incorporate specific monitoring tools, such as interactive dashboards and decision-support systems. These tools become pivotal in providing a comprehensive visualization of the results, enabling physicians and AI experts to assess the models' performance.

4.2. The role of MLOps

During our experience inside the DARE foundation, we encountered several challenges related to the development environment. To face them and be compliant with existing regulations, we thought that MLOps was the best solution to ensure high-quality data and release safe models. For instance, thanks to its CI/CD nature, MLOps can monitor the related outcomes by providing accurate prediction tools for the diagnostic and prognostic processes [12].

Therefore, starting from the healthcare pipeline shown in Fig. 3, we continue its definition by explaining the role played by MLOps. To this end, we first describe the main MLOps steps. Then, we conceptually map such steps over our pipeline and show how MLOps fully supports the entire workflow. According to the definition provided by Moskalenko et al. [17], the MLOps development cycle, however complex it is, can be summarized in the following steps shown in Fig. 4:

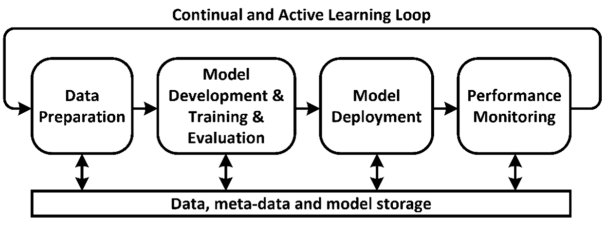

Figure 4: The development cycle of MLOps pipelines [17].

1. **Data Preparation:** it is responsible for data management and ensures that the employed datasets comply with all healthcare standards and regulations [8, 34]. To this end, several tools, such as DVC [35], Deepchecks [36], and Great Expectations [37], can support developers in collecting, gathering, cleaning, and transforming data before the training process;
2. **Model Development:** it iteratively builds models by tracking all related hyperparameters. Typically, thanks to their large communities, MLflow [38] and ClearML [39] are two of the most famous platforms employed to manage this step;
3. **Model Deployment:** it integrates the developed models into specific services or diagnostic tools. For this purpose, MLOps frameworks rely on CI/CD pipelines (e.g., built by combining the tools mentioned above) to deploy ML models into production environments [14];
4. **Performance Monitoring:** it monitors the reliability of the deployed models. During this step, tools like Prometheus [40] and Grafana [41] can provide visual dashboards to detect model drift through the analysis of the related inputs, outputs, and metrics [12].

According to the given definitions, these steps are overlappable with those shown in Fig. 3. More precisely, Data Preparation covers the same roles as Data Collection. They are responsible for data collection, cleaning, and transformation that may come from different healthcare scenarios. Follows the Model Development step, which covers some functions of Data Processing and Model Management. Thanks to the most famous MLOps platforms (i.e., MLflow [38] and ClearML [39]), Model Development ensures that the collected data are continuously stored and monitored. Similarly, the Model Deployment step also covers some functions of Data Processing and Model Management. In detail, it rigorously manages models in production by following a CI/CD logic. For this reason, the stored data do not represent only those coming from patients but also those related to models (e.g., provided outputs, considered hyperparameters, and tracked metrics). Finally, the Performance Monitoring and Results Visualization steps aim to monitor the models' drift and data quality through dedicated graphical interfaces, as well as the data related to new patients (i.e., those not considered during the training process). Ultimately, adopting an MLOps framework not only supports the development of a healthcare pipeline but also provides additional benefits for the involved models.

5. Ensure the CSF requirements

This section aims to examine the feasibility of MLOps in ensuring CSF requirements. To this end, starting from the defined pipeline, we show how MLOps can enable one of the most important Categories of the Protect (PR) Function, namely Data Security. Then, we also highlight that our pipeline ensures other CSF Categories, focusing on those related to Detect (DT) and Recover (RC) Functions. For each Category, we report the related results through checklist tables containing the Subcategory ID (ID), the Subcategory definition (Definition), and the ensure type (i.e., direct, indirect, and no). Notice that we derived such tables by faithfully following those reported by CSF [23]. Furthermore, we use the term direct for those requirements that MLOps ensures, together with best practices, without further implementation steps. Instead, with indirect, we refer to all those requirements in which MLOps needs to interface with additional implementations or unrelated tools.

5.1. MLOps in ensuring Data Security

With reference to Fig. 2, we start our discussion by showing how MLOps can ensure Data Security (DS). To this end, we assume that our pipeline complies with all cybersecurity best practices (i.e., effectively implements and employs communication protocols, access controls, and encryption techniques) [19, 42]. Inadequate data protection measures can lead to unauthorized access to sensitive information, resulting in potential legal and financial consequences. Not surprisingly, one of the primary concerns is represented by data breaches, which can occur in different stages [43]. For this reason, the following assumption also provides a fertile ground for MLOps to collect, process, and store data securely. From a practical point of view, this means employing the CI/CD nature to encrypt sensitive data during all

lifecycle (i.e., at rest, in transit, and in use) and conduct regular audits to monitor compliance with data protection regulations [43].

For this reason, as shown in Tab. 2, MLOps allows the consistent management of data with risk management strategies [23], by directly ensuring confidentiality, integrity, and availability throughout the development cycle (see PR.DS-01, PR.DS-02, and PR.DS-10). As discussed in Sec. 4, the role played by MLOps becomes more pivotal because it forces organizations to adopt a multi-dimensional approach that includes a range of advanced techniques and tools [43], such as MLFlow [38], DVC [35], and Great Expectations [37]. All this is possible because these tools give MLOps pipelines the fundamental ability to track any aspect (i.e., data, hyperparameters, metrics, and models) [44], minimize performance degradation, and manage data drift [13]. As a result, the following ability also makes the creation of uniquely identified and rigorously maintained backups (see PR.DS-11). Therefore, we can state that the application of MLOps within our pipeline can directly ensure the DS requirements reported in Tab. 2.

ID	Definition	Ensure type
PR.DS-01	The confidentiality, integrity, and availability of data-at-rest are protected	direct
PR.DS-02	The confidentiality, integrity, and availability of data-in-transit are protected	direct
PR.DS-10	The confidentiality, integrity, and availability of data-in-use are protected	direct
PR.DS-11	Backups of data are created, protected, maintained, and tested	direct

Table 2
Requirements related to Data Security (DS).

5.2. MLOps compared to other CSF Functions

Subsequently, by adopting the same methodology, we also focused on the remaining Functions (i.e., Govern - GV, Identify - ID, Detect - DT, Respond - RS, and Recover - RC). With reference to Fig. 2, the results of this iterative process have highlighted interesting evidence for some Categories of DT and RC Functions. Concerning DT, we have found some correspondences in the Continuous Monitoring (CM) Category. As shown in Tab. 3, CM defines some requirements for ensuring that assets are monitored to find anomalies, indicators of compromise, and other potentially adverse events [23]. Among the requirements identified with DE.CM, only two are indirectly ensured: DE.CM-01 and DE.CM-09. More precisely, the ability to monitor models in production allows to count the number of interactions made. For instance, during a Denial-of-Service (DoS) attack, there could be numerous "unnecessary" requests aimed at saturating the responsiveness of the hosting asset. Therefore, by employing dedicated user interfaces, such as those implemented with Prometheus [40] and Grafana [41], it is possible to monitor networks and the related services (see DE.CM-01) by considering the number of requests, the received inputs, and the provided outputs. Consequently, together with the ability to record any experimental aspect, this also allows MLOps to monitor runtime environments and the employed data (see DE.CM-09).

Instead, for the RC Function, we have found some correspondences in the Incident Recovery Plan Execution (RP) Category. As shown in Tab. 4, RP defines some requirements for the correct restoration activities, which are performed to ensure the operational availability of systems and services affected by cybersecurity incidents [23]. Among the requirements identified with RC.RP, only three are indirectly ensured: RC.RP-01, RC.RP-02, and RC.RP-03. More precisely, thanks again to its ability to record any aspect of the development lifecycle, MLOps indirectly generates backups of the employed dataset and models. Such backups can be useful when the recovery portion of the incident response plan is executed (see RC.RP-01). Moreover, this ability also supports the selection, prioritization, and execution of recovery actions (see RC.RP-02). For example, running a pipeline's step rather than or before another. Finally, with the support of some notable technologies (e.g., DVC [35], Deepchecks [36], and Great Expectations [37]), MLOps can verify the backup integrity before the recovery (see RC.RP-03).

ID	Definition	Ensure type
DE.CM-01	Networks and network services are monitored to find potentially adverse events	indirect
DE.CM-02	The physical environment is monitored to find potentially adverse events	no
DE.CM-03	Personnel activity and technology usage are monitored to find potentially adverse events	no
DE.CM-06	External service provider activities and services are monitored to find potentially adverse events	no
DE.CM-09	Computing hardware and software, runtime environments, and their data are monitored to find potentially adverse events	indirect

Table 3
Requirements related to Continuous Monitoring (CM).

ID	Definition	Ensure type
RC.RP-01	The recovery portion of the incident response plan is executed once initiated from the incident response process	indirect
RC.RP-02	Recovery actions are selected, scoped, prioritized, and performed	indirect
RC.RP-03	The integrity of backups and other restoration assets is verified before using them for restoration	indirect
RC.RP-04	Critical mission functions and cybersecurity risk management are considered to establish post-incident operational norms	no
RC.RP-05	The integrity of restored assets is verified, systems and services are restored, and normal operating status is confirmed	no
RC.RP-06	The end of incident recovery is declared based on criteria, and incident-related documentation is completed	no

Table 4
Requirements related to Incident Recovery Plan Execution (RP).

Despite the contribution highlighted in this work, it appears that the proposed MLOps pipeline ensures only a few requirements (i.e., Subcategories). However, according to the definition provided by CSF, this aspect does not necessarily represent a limitation. Regardless of the size or importance of the organization concerned, the CSF should be used in conjunction with other resources (e.g., frameworks, standards, guidelines, and leading practices) to manage cybersecurity risks as well as possible [23]. Moreover, from a practical point of view, it is impossible to cover all cybersecurity aspects by employing only one technology [45]. Therefore, on the basis of the discussed outcomes, we consider essential to investigate the remaining CSF Functions (i.e., those not covered) by considering different scenarios or enhanced frameworks (e.g., the Machine Learning Security Operations [46]). Finally, the feasibility of MLOps pipelines should be evaluated with respect to legal requirements, such as those defined by the AI Act [47] and European Regulation (2017/745) [48].

6. Conclusions and Future Work

Implementing Machine Learning (ML) models for healthcare scenarios represents a challenging activity, ranging from data quality management to compliance with stringent regulations. In this context, MLOps pipelines emerge as promising solutions for managing the lifecycle of developed models, which is vital for diagnostic and prognostic activities. On the other hand, the development of healthcare systems should also consider several cybersecurity aspects strictly related to such regulations. In response to these additional challenges, the Cybersecurity Framework (CSF) 2.0, defined by the National Institute of Standards and Technology (NIST), provides updated guidelines to address security issues in an ever-evolving technological landscape. For this reason, we investigated the feasibility of MLOps pipelines in ensuring the requirements defined by CSF. To this end, we first presented an overview of the fundamental concepts employed, namely the CSF-related structure (i.e., Functions and Categories) and the main characteristics of MLOps. Then, based on our experience with the DARE foundation, we presented the high-level architecture of a healthcare MLOps pipeline. Finally, by adopting the CSF, we discussed the feasibility of our pipeline in ensuring Data Security, which represents one of the most important Categories of the Protect (PR) Function. Moreover, by iteratively analyzing the remaining CSF Functions, we have also highlighted that MLOps might indirectly ensure other CSF Categories, with particular emphasis on those of Detect (DT) and Recover (RC).

However, due to the numerous, heterogeneous, and high-level requirements defined in CSF, it is impossible to cover all related aspects in the following study. For this reason, we will investigate MLOps pipelines and their benefits by considering other healthcare scenarios. This first contribution will allow us to analyze the remaining CSF Functions. Moreover, to improve the achieved outcomes, we will also combine enhanced frameworks, such as Machine Learning Security Operations (MLSecOps), with the implementations of real use cases. Finally, since we presented a pipeline employed by a real healthcare foundation, we will also analyze the feasibility of MLOps in ensuring legal requirements, such as those defined by the AI Act and European Regulation (2017/745).

Acknowledgments

This study was partially supported by the Italian Ministry of University and Research under PNRR-PNC Project PNC0000002 “DARE—Digital Lifelong Prevention” (CUP: B53C22006450001).

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT, Grammarly in order to: Grammar and spelling check, Paraphrase and reword. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the publication’s content.

References

[1] M. M. Ahsan, S. A. Luna, Z. Siddique, Machine-learning-based disease diagnosis: A comprehensive review, Healthcare (2022). doi:10.3390/healthcare10030541.

[2] G. Battineni, G. G. Sagaro, N. Chinatalapudi, F. Amenta, Applications of machine learning predictive models in the chronic disease diagnosis (2020). doi:10.3390/jpm10020021.

[3] D. Bertsimas, A. Orfanoudaki, R. B. Weiner, Personalized treatment for coronary artery disease patients: a machine learning approach (2020). doi:10.1007/s10729-020-09522-4.

[4] B. Schwartz, Z. D. Cohen, J. A. Rubel, D. Zimmermann, W. W. Wittmann, W. Lutz, Personalized treatment selection in routine care: Integrating machine learning and statistical algorithms to recommend cognitive behavioral or psychodynamic therapy, Psychotherapy Research (2020).

[5] P. N. Ramkumar, H. S. Haeberle, D. Ramanathan, W. A. Cantrell, S. M. Navarro, M. A. Mont, M. Bloomfield, B. M. Patterson, Remote patient monitoring using mobile health for total knee arthroplasty: Validation of a wearable and machine learning–based surveillance platform, The Journal of Arthroplasty 34 (2019) 2253–2259. doi:10.1016/j.arth.2019.05.021.

[6] A. Rghioui, J. Lloret, S. Sendra, A. Oumnad, A smart architecture for diabetic patient monitoring using machine learning algorithms, Healthcare 8 (2020). doi:10.3390/healthcare8030348.

[7] L. E. Lwakatare, A. Raj, I. Crnkovic, J. Bosch, H. H. Olsson, Large-scale machine learning systems in real-world industrial settings: A review of challenges and solutions, Information and Software Technology (2020). doi:10.1016/j.infsof.2020.106368.

[8] E. Petersen, Y. Potdevin, E. Mohammadi, S. Zidowitz, S. Breyer, D. Nowotka, S. Henn, L. Pechmann, M. Leucker, P. Rostalski, C. Herzog, Responsible and regulatory conform machine learning for medicine: A survey of challenges and solutions (2022). doi:10.1109/ACCESS.2022.3178382.

[9] S. Vänskä, K.-K. Kemell, T. Mikkonen, P. Abrahamsson, Continuous software engineering practices in AI/ML development past the narrow lens of MLOps: Adoption challenges, e-Informatica Software Engineering Journal 18 (2024) 240102. doi:10.37190/e-inf240102.

[10] F. Calefato, L. Quaranta, F. Lanubile, M. Kalinowski, Assessing the use of automl for data-driven software engineering, 2023 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (2023) 1–12.

[11] F. Lanubile, F. Calefato, L. Quaranta, M. Amoruso, F. Fumarola, M. Filannino, Towards productizing ai/ml models: An industry perspective from data scientists, in: 2021 IEEE/ACM 1st Workshop on AI Engineering (WAIN), 2021. doi:10.1109/WAIN52551.2021.00027.

[12] M. M. John, H. H. Olsson, J. Bosch, Towards mlops: A framework and maturity model, in: 2021 47th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), 2021.

[13] B. M. A. Matsui, D. H. Goya, Mlops: Five steps to guide its effective implementation, in: 2022 1st International Conference on AI Engineering (CAIN), 2022. doi:10.1145/3522664.3528611.

[14] I. Karamitsos, S. Albarhami, C. Apostolopoulos, Applying devops practices of continuous automation for machine learning, Information 11 (2020). doi:10.3390/info11070363.

[15] P. S. U. Shah, N. Ahmad, M. O. Beg, Towards mlops: A devops tools recommender system for machine learning system (2024). URL: https://api.semanticscholar.org/CorpusID:267759567.

[16] T. Mboweni, T. Masombuka, C. Dongmo, A systematic review of machine learning devops, in: 2022 International Conference on Electrical, Computer and Energy Technologies (ICECET), 2022.

[17] V. Moskalenko, V. Kharchenko, Resilience-aware mlops for ai-based medical diagnostic system, Frontiers in Public Health 12 (2024). doi:10.3389/fpubh.2024.1342937.

[18] M. Reddy, B. Dattaprakash, S. Kammath, S. Kn, S. Manokaran, R. Be, Application of mlops in prediction of lifestyle diseases, ECS Transactions 107 (2022) 1191. doi:10.1149/10701.1191ecst.

[19] T. Ahmad, M. Adnan, S. Rafi, M. A. Akbar, A. Anwar, Mlops-enabled security strategies for next-generation operational technologies, in: Proc. of the 28th International Conference on Evaluation and Assessment in Software Engineering, EASE '24, 2024. doi:10.1145/3661167.3661283.

[20] M. Fredrikson, S. Jha, T. Ristenpart, Model inversion attacks that exploit confidence information and basic countermeasures, in: Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, CCS '15, 2015. doi:10.1145/2810103.2813677.

[21] I. Rosenberg, A. Shabtai, Y. Elovici, L. Rokach, Adversarial machine learning attacks and defense methods in the cyber security domain, ACM Comput. Surv. 54 (2021). doi:10.1145/3453158.
[22] F. Calefato, F. Lanubile, L. Quaranta, Security risks and best practices of MLOps: A multivocal literature review, in: ITASEC 2024: Italian Conference on Cybersecurity, Salerno, Italy, 2024.
[23] The NIST Cybersecurity Framework (CSF) 2.0, 2024. doi:10.6028/nist.cswp.29.
[24] DARE Foundation, 2024. URL: https://www.fondazionedare.it/en/.
[25] F. Calefato, F. Lanubile, L. Quaranta, A preliminary investigation of mlops practices in github, in: Proceedings of the 16th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement, ESEM '22, 2022. doi:10.1145/3544902.3546636.
[26] H. Villamizar, M. Kalinowski, H. Lopes, D. Mendez, Identifying concerns when specifying machine learning-enabled systems: A perspective-based approach, Journal of Systems and Software (2024).
[27] A. Basile, F. Calefato, F. Lanubile, G. Mallardi, L. Quaranta, An mlops solution framework for transitioning machine learning models into ehealth systems, 2024, pp. 318–323.
[28] A. Krishnan, V. Subasri, K. McKeen, A. Kore, F. Ogidi, M. Alinoori, N. Lalani, A. Dhalla, A. Verma, F. Razak, D. Pandya, E. Dolatabadi, Cyclops: Cyclical development towards operationalizing ml models for health, medRxiv (2022). doi:10.1101/2022.12.02.22283021.
[29] G. Danciu, I. E. Nicolae, I. Ilie, C. S. Nechifor, Advanced notebook: A tool for enhanced management of machine learning models and procedures in the healthcare domain, in: International Conference on Applied Mathematics & Computer Science (ICAMCS), 2023, pp. 36–41.
[30] T. Granlund, V. Stirbu, T. Mikkonen, Towards regulatory-compliant mlops: Oravizio's journey from a machine learning experiment to a deployed certified medical product (2021).
[31] S. Garg, P. Pundir, G. Rathee, P. Gupta, S. Garg, S. Ahlawat, On continuous integration / continuous delivery for automated deployment of machine learning models using mlops, in: 2021 IEEE Fourth International Conference on Artificial Intelligence and Knowledge Engineering (AIKE), 2021.
[32] B. Ghosh, Adopting mlsecops: Securing machine learning at scale, 2023. URL: https://medium.com/@bijit211987/adopting-mlsecops-securing-machine-learning-at-scale-1a5647d01a64.
[33] D. Kreuzberger, N. Kühl, S. Hirschl, Machine learning operations (mlops): Overview, definition, and architecture, IEEE Access 11 (2023) 31866–31879. doi:10.1109/ACCESS.2023.3262138.
[34] O. Diaz, K. Kushibar, R. Osuala, A. Linardos, L. Garrucho, L. Igual, P. Radeva, F. Prior, P. Gkontra, K. Lekadir, Data preparation for artificial intelligence in medical imaging: A comprehensive guide to open-access platforms and tools (2021). doi:10.1016/j.ejmp.2021.02.007.
[35] DVC: Data Version Control, 2024. URL: https://dvc.org/.
[36] Deepchecks, 2021. URL: https://github.com/deepchecks/deepchecks.
[37] Great Expectations: Always know what to expect, 2024. URL: https://greatexpectations.io/.
[38] MLflow: Open-source platform for the machine learning lifecycle, 2024. URL: https://mlflow.org/.
[39] Clearml - your entire mlops stack in one open-source tool, 2024. URL: https://clear.ml/.
[40] Prometheus: Monitoring system and time series database, 2024. URL: https://prometheus.io/.
[41] Grafana: Open-source analytics and monitoring solution, 2024. URL: https://grafana.com/.
[42] A. Lima, L. Monteiro, A. P. Furtado, Mlops: Practices, maturity models, roles, tools, and challenges - A systematic literature review, 2022. doi:10.5220/0010997300003179.
[43] A. Thompson, Enhancing model security in devops pipelines: A comprehensive approach to mlops security, Distributed Learning and Broad Applications in Scientific Research 10 (2024) 332–338.
[44] H. Safri, G. Papadimitriou, E. Deelman, Dynamic tracking, mlops, and workflow integration: Enabling transparent reproducibility in machine learning, in: 2024 IEEE 20th International Conference on e-Science (e-Science), 2024, pp. 1–10. doi:10.1109/e-Science62913.2024.10678658.
[45] M. Lezzi, M. Lazoi, A. Corallo, Cybersecurity for industry 4.0 in the current literature: A reference framework (2018). doi:https://doi.org/10.1016/j.compind.2018.09.004.
[46] A. Saputra, E. Suryani, N. A. Rakhmawati, The robustness of machine learning models using mlsecops: A case study on delivery service forecasting, in: 2023 14th International Conference on Information & Communication Technology and System (ICTS), 2023, pp. 265–270.
[47] The EU Artificial Intelligence Act, https://artificialintelligenceact.eu/the-act/, 2024.
[48] Medical Device Regulation - 2017/745, http://data.europa.eu/eli/reg/2017/745/oj, 2017.